1.0 Introduction

The initial phase of the Columbia River Comprehensive Impact Assessment (CRCIA) is a screening assessment of risk to humans and the environment. To assess risk, monitoring data of contaminant concentrations are needed. The data task provides measurements of contaminant concentrations in various media for use in the human health and ecological screening assessments. This report is divided into two volumes. *Volume I: Text* describes the data gathering and data selection processes. *Volume II: Appendices* presents the 1) final data sets (media files) to be used in the screening assessments, 2) the raw data from which the data sets were derived, and 3) the raw data values for media for which other calculation methods will be used. (For a copy of Volume II with 500 pages plus 9 diskettes, contact S. D. Cannon at 509-372-6210.)

1.1 Scope

The Columbia River has been the focus of environmental monitoring programs over the years. The scope of the data task is to compile data collected since January 1990 by the various monitoring programs for the contaminants of interest. The contaminants of interest for the screening assessment were originally defined in Napier et al. (1995) and have been revised based on comments received on that draft document. The revised list of contaminants is given in Table 1.1.

Table 1.1. List of Identified Contaminants of Interest[a]

Radionuclides	Metals	Other Compounds
Carbon-14	Chromium	Ammonia
Cesium-137	Copper	Benzene
Cobalt-60	Lead	Cyanide
Europium-152	Mercury	Diesel oil
Europium-154	Nickel	Kerosene
Iodine-129	Zinc	Nitrate
Neptunium-237		Nitrite
Strontium-90		Phosphates
Technetium-99		Sulfates
Tritium		Xylenes
Uranium-234		
Uranium-238		
(a) Direct irradiation and discrete radioactive particles will also be evaluated.		

The media for which concentration data are needed for the human health and ecological screening assessment calculations are groundwater, sediment, seeps, surface water, and external radiation. These media files along with the original raw data files are presented in this report. In addition, contaminant concentrations in biota, cobalt-60 particles, drive point groundwater data for chromium, N Springs punch point water data, and pore water data for chromium will be evaluated in the screening assessment. These raw data values are also presented in this report. However, because the availability of data applicable to the screening assessment is limited, other calculation methods will be used in the screening assessment for biota, cobalt-60 particles, drive point groundwater, N Springs punch point water, and pore water. Therefore, no media files needed to be prepared for these data.

1.2 Approach

The data task for CRCIA is being conducted under the guidance of the CRCIA Management Team. All defining decisions for the task were made with CRCIA Team concurrence. All team decisions relating to the efforts of the data task are presented in Table 1.2.

Table 1.2. Data Decisions by the CRCIA Team

Date	Decision
1/30/96	Agreement was reached to collect data from January 1, 1990 to present and fill data gaps with older data where it is available for the initial phase of the screening assessment.
1/30/96	The primary geographic focus area for the screening assessment is from Priest Rapids Dam to McNary Dam. A radionale will be provided justifying this area by including in the report a discussion of historical levels/trends in contaminant data over time showing levels typically upstream of McNary, including Hanford data, Oregon data, and Washington data were available.
2/13/96	All data will be provided on a diskette in the final report.
2/13/96	There will be no soil medium. There are no soil samples associated with the outfall pipe locations, and no other soil data were requested by the human health and ecological task leads.
2/13/96	The river (between Priest Rapids and McNary) will be broken into 27 segments. This partially defines the spatial aggregation of the data.
2/13/96	Corridor widths were chosen by segment based on sampling sites available to characterize contamination. Reactor areas 100 B/C, D, F, H, K, N and the 300 Area have 0.4-kilometer (1/4-mile) corridor widths. (N Reactor width was changed from 0.8 kilometer to 0.4 kilometer at 3/5/96 CRCIA Team meeting.) The non-trench portion of the 100-K Area has a 0.6-kilometer (3/8-mile) corridor width. All other segments have a 0.8-kilometer (1/2-mile) width. This completes the definition of the spatial aggregation of the data.
2/13/96	A representative value for each groundwater well in each segment will be chosen. A mechanized process needs to be developed to choose the representative value. It is expected that the mechanized process will be adequate for about 80 percent of the values. Remaining values will need to be looked at by hand. A team was formed to develop the algorithm.
2/20/96	Where there is a clear upward or downward trend, a representative value will be chosen from the most recent data.
2/20/96	The maximum representative value for each data set should be an observed data point.

Table 1.2. (contd)

Date	Decision
2/20/96	The set of representative data in each segment for each medium will be assumed to be lognormally distributed. The parameters for the lognormal distribution will be estimated from the representative data. Log-probability plots will be provided.
2/20/96	Both filtered and unfiltered data will be used in the identification of representative data and in determining the parameters for the lognormal distribution.
2/27/96	Dixon's test will be used to eliminate, at most, a single outlier data point in each data set. In the data section of the final report, every data point that is eliminated will be explained.
2/27/96	For the elimination of outliers, log transformation of the data will be used.

A Geographic Information System was used to assist in implementing the processing of the data for the screening assessment. The Geographic Information System is a computerized system designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information. Many software packages exist that perform basic Geographic Information System functions. Arc/Info Rev. 7.0.2 was used for the data task (ESRI 1994).

1.2.1 Segmentation

The human health and ecological screening assessments calculate risk based on contributions from multiple pathways affected by contaminant concentrations in multiple media. These contaminant concentrations were not usually measured in a fashion that would allow a complete assessment at every sampling site. To provide data for the assessments, it was necessary to aggregate data to represent concentrations in areas rather than at points. This was done through the technique of river segmentation.

Staff from the Pacific Northwest National Laboratory (PNNL), U.S. Environmental Protection Agency (EPA), and Washington State Department of Ecology (Ecology) defined 27 segments within the study area from Priest Rapids Dam down to McNary Dam (see Figure 1.1). A segment is a section of the river over which contaminant conditions can be expected to be similar and which captures the major influences to the Columbia River. The major resources used to decide how to most appropriately segment the river were a groundwater well location map (Dresel et al. 1995), the radiological and chemical contaminant plume maps from the 1994 Hanford Site groundwater monitoring report (Dresel et al. 1995), the Surface Environmental Surveillance Project spring monitoring locations information, the maps of sampling locations for the special chromium studies being conducted by the Environmental Restoration Contractor at the 100-H Area (Hope and Peterson 1995) and 100-D Area (from an unpublished document by S. J. Hope and R. E. Peterson), and the Ecology and EPA staff knowledge of the contaminant sources.

Because many contaminant sources are located in the reactor areas, each reactor area was examined to determine if a single segment could be designed around it or whether it should be further divided. When this decision was made, the resources listed above were used to determine the upstream
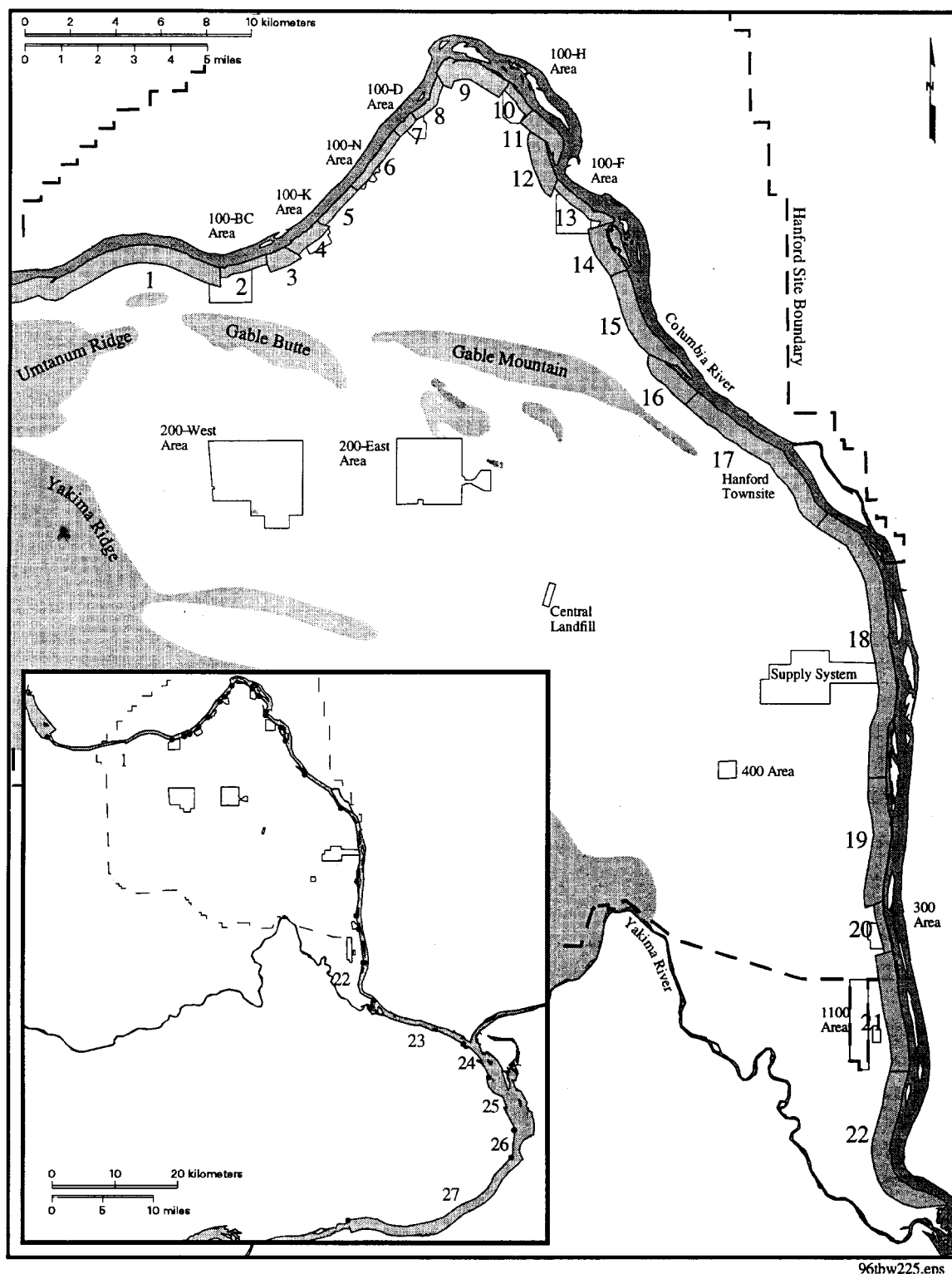
Figure 1.1. Segmentation of the Columbia River

and downstream cutoff points for these segments. This process was continued for other major features of the river, such as the sloughs and the confluence points of the Yakima, Snake, and Walla Walla Rivers. The final step in the segmentation process was to look at the areas between known areas of interest and determine how to define segments in these generally data sparse areas.

1.2.2 Thiessen Polygon Analysis

A Thiessen polygon analysis was used to define the area of influence of the data from a ground-water well and, thereby, refine the segmentation of the river. Thiessen analyses apportion points into polygonal regions such that each region contains only one point. Each region has the unique property that any location within a region is closer to the region's point than to any other point (Thiessen and Alter 1911). In this study, the points used for the analysis are the groundwater well locations. An example of the results of the Thiessen analysis are shown in Figure 1.2. The example shows a portion of the study area where segment boundaries were adjusted based on the Thiessen results.

Adjustments to segment boundaries were based on the Thiessen analysis for two reasons. First, it was desirable for the polygons within a segment to be of similar size so that each well is used to represent the contaminant concentration over an area of similar size. When polygon sizes became large relative to the other polygons in the segment, then the large polygon was clipped and the space was added to neighboring segments. This phenomenon occurs with the outer most wells in a data dense area when the data dense area borders an area with few or no wells. Therefore, the additional space was added to a data sparse segment. Second, the original segment boundaries were drawn perpendicular to the river shore. To better reflect the areas of well influence, the polygon lines were used to represent the segment boundaries. Examples of this type of adjustment can be seen on Figure 1.2 for both the upstream and downstream boundaries of segment 3.

1.2.3 Corridor

Data for all media were initially gathered from a corridor up to 0.8 kilometer (½ mile) on either side of the Columbia River. For sediment, seeps, surface water, and external radiation, all data within 0.8 kilometer of the river were used. For the groundwater data, it was necessary to use only the portion of these data that would be relevant to estimating the contaminant concentrations entering the Columbia River from the Hanford Site. This was done by assigning a groundwater corridor width to the Hanford side of each segment. The corridor width was based on having sufficient groundwater data to characterize the contamination within a segment. These corridor width decisions were made by staff from Ecology, EPA, and PNNL with concurrence by the CRCIA Team. The corridor widths for groundwater data are as follows. Segments 2 (100-B/C Area), 5 (116-K-2 trench in 100-K Area), 6 (100-N Area), 7 and 8 (100-D Area), 10 (100-H Area), 13 (100-F Area), and 20 (300 Area) have a 0.8-kilometer (1/4-mile) corridor width. Segment 4 (100-K Area) has a 0.6-kilometer (3/8-mile) corridor width. All other segments have a 0.8-kilometer (1/2-mile) width (Figure 1.1). These ground-water corridor widths are measured from the Columbia River shoreline inland onto the Hanford Site.
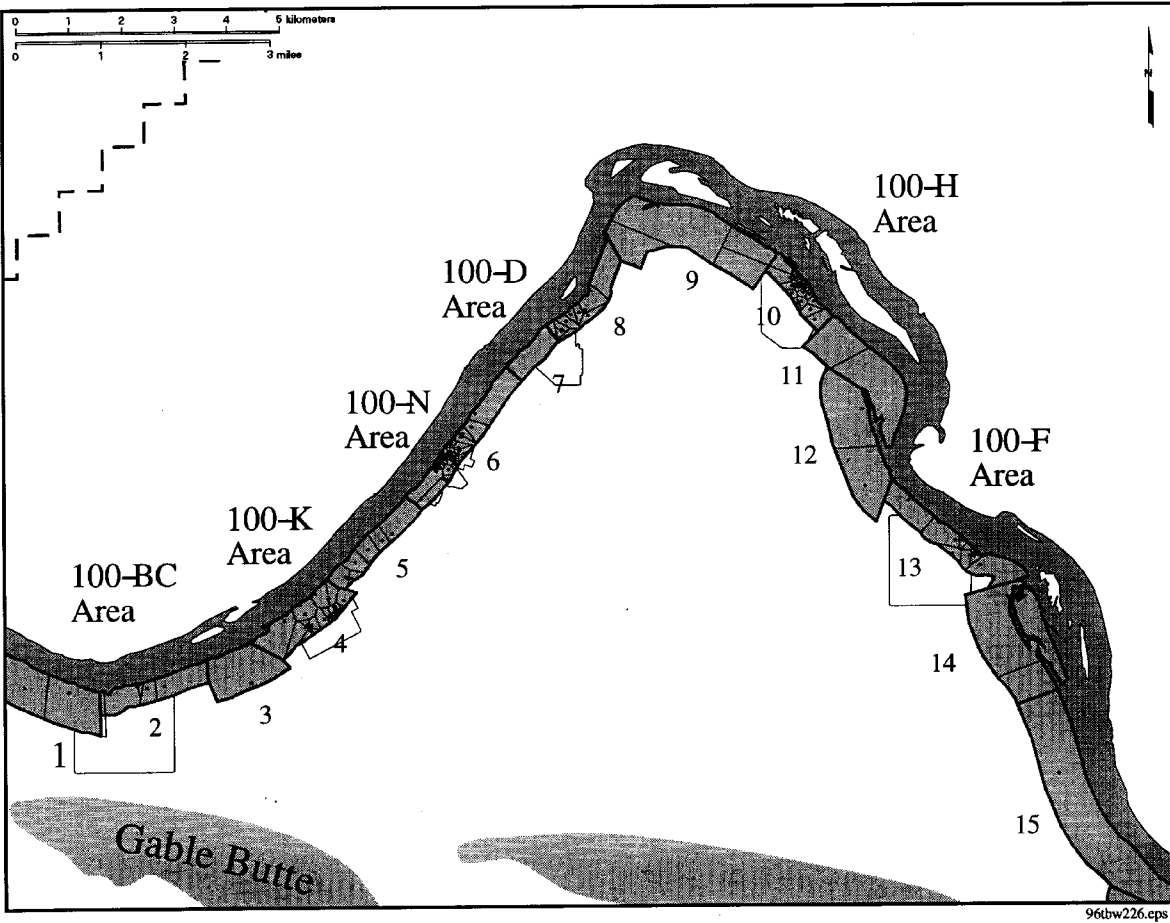
Figure 1.2. Example of the Results of the Thiessen Analysis on Segmentation of the Columbia River